

- 1 -

Pre-Allocating Resources Of A Wireless Network
For Packet-Switched Real-Time, Interactive Communications

Technical Field

5 [0001] The invention relates generally to pre-allocating resources of a wireless network for packet-switched real-time, interactive communications.

Background

10 [0002] A mobile communications network is typically made up of a plurality of cells. Each cell includes a radio base station, with each base station connected to a mobile switching center or a packet service node that manages communications sessions between mobile stations and terminals coupled to a public switched telephone network (PSTN) or a packet-based data network. Communications between mobile stations and base stations are performed over wireless links.

15 [0003] Traditional wireless protocols provide for circuit-switched communications. Such protocols include time-division multiple access (TDMA) protocols and code-division multiple access (CDMA) protocols. In a circuit-switched network, a channel portion between two endpoints (*e.g.*, two mobile stations) is occupied for the duration of the connection between the endpoints.

20 [0004] With the wide availability of the Internet and intranets, packet-switched communications (*e.g.*, web browsing, electronic mail, and so forth) have become more common. Generally, a circuit-switched connection is an inefficient mechanism for communicating packet data. As a result, third generation (3G) and beyond wireless technologies are being developed and implemented to provide higher bandwidth and more efficient packet-switched communications (of data as well as voice and other forms of
25 real-time data) over wireless networks.

[0005] One example of a packet-switched wireless technology is defined by the CDMA2000 family of standards, developed by the Third Generation Partnership Project 2 (3GPP2). A CDMA2000 wireless communications network is capable of supporting both circuit-switched services and packet-switched services. For TDMA, packet-switched wireless
30 communications protocols have also been developed, such as the Enhanced General Packet Radio Service (EGPRS) protocol as defined by the 3GPP (Third Generation Partnership

- 2 -

Project) UMTS (Universal Mobile Telecommunications System) Release 1999 Standard, and others.

[0006] One form of communicating of voice in packet-switched communications is referred to as voice-over-Internet Protocol (IP). In voice-over-IP, voice (and other forms of real-time data) is carried in IP packets in an IP session established between two or more network devices. With advancements in packet-switched wireless technologies, voice-over-IP over packet-switched wireless networks have also been implemented.

[0007] A more recent advancement is the proposal of press (push)-to-talk (PTT) over voice-over-IP in a wireless network. This technology is based on PTT over cellular (PoC) technology, which enables real-time, one-to-one or one-to-many voice communications service over a wireless network that is started by pressing or pushing a talk key or button on a mobile station. PTT enables multiple users to communicate with each other, where one party (the caller) has control and right-to-speak at any one time. To acquire the right-to-speak, the caller sends a request (referred to as a floor control request) to a PTT server, where the request is sent in response to pushing of the talk key or button on a mobile station.

[0008] To set up a packet-switched real-time, interactive communications session, certain resources are allocated during a call setup procedure for the communications session. The resources allocated during call setup include various software, hardware, and communications resources in nodes of the wireless network as well as in other nodes (e.g., a PTT server or other type of server that supports packet-switched real-time, interactive communications). In many cases, the allocation of resources during the call setup procedure is relatively time consuming, which adds to the overall delay experienced by the user in setting up a packet-switched real-time, interactive communications session. The delay may be unacceptable or inconvenient for certain delay-sensitive applications.

Summary

[0009] In general, according to an embodiment, a method of communicating in a wireless network includes pre-allocating, to a packet-switched real-time, interactive communications application, resources of at least one node of the wireless communications network. The pre-allocated resources comprising resources normally allocated in response to a call setup request. A first call setup request is received after pre-allocating the resources, and, in

- 3 -

response to the first call setup request, a packet-switched real-time, interactive communications session is established through the wireless network using the pre-allocated resources of the at least one node.

[0010] Other or alternative features will become apparent from the following description,
5 from the drawings, and from the claims.

Brief Description Of The Drawings

[0011] Fig. 1 is a block diagram of an example communications network that incorporates an embodiment.

10 [0012] Fig. 2 is a block diagram of another example communications network that incorporates an embodiment.

[0013] Fig. 3 is a message flow diagram of a process of provisioning nodes of a communications network and setting up a packet-switched real-time, interactive communications session, in accordance to an embodiment.

Detailed Description

15 [0014] In the following description, numerous details are set forth to provide an understanding of some embodiments. However, it will be understood by those skilled in the art that some embodiments may be practiced without these details and that numerous variations or modifications from the described embodiments may be possible.

20 [0015] Referring to Fig. 1, an example communications network includes a wireless or mobile communications network that is coupled to a packet data network 34 (such as a local area network (LAN), wide area network (WAN), Internet, and so forth). According to an embodiment, the wireless communications network includes components that operate according to CDMA (code-division multiple access) 2000. CDMA2000 is defined by the CDMA2000 family of standards (including the TIA-2000 standards, TIA-2001 standards, and
25 the TIA-835 standards). However, in other embodiments, other types of wireless protocols can be used for communications in the wireless communications network, including other versions of CDMA protocols, TDMA (time-division multiple access) protocols, GSM

- 4 -

(Global Systems for Mobile Communications), UMTS (Universal Mobile Telecommunications System) protocols, or other wireless protocols.

[0016] The wireless communications network includes multiple cell segments each including a base transceiver subsystem (BTS) 20A, 20B for performing radio telecommunications with mobile stations 17A, 17B within the coverage area of the respective cell segment. A “cell segment” refers to either a cell or cell sector. The BTS entities 20A, 20B are connected to one or more base station controllers (BSCs) 22.

[0017] For communicating circuit-switched voice traffic, the BSC 22 is coupled to a mobile switching center (MSC) 24, which is responsible for switching mobile station-originated or mobile station-terminated circuit-switched traffic. Effectively, the MSC 24 is the interface for signaling and user traffic between the wireless network and other circuit-switched networks (such as a public switched telephone network (PSTN) 26 or other MSCs). The PSTN 26 is connected to landline terminals (not shown).

[0018] The wireless communications network also supports packet data services, in which packet data is communicated between a mobile station and another endpoint, which can be a terminal coupled to the packet data network 34 or another mobile station that is capable of communicating packet data. Packet data is communicated in a packet-switched communications session established between the mobile station and the other endpoint through a packet data serving node (PDSN) 30. For other types of wireless networks, the PDSN 30 can be substituted with an SGSN (Serving GPRS support node), a GGSN (Gateway GPRS support node), or an IWF (inter-working function) server.

[0019] The PDSN 30 establishes, maintains, and terminates link layer sessions to mobile stations, and routes mobile station-originated or mobile station-terminated packet data traffic. The PDSN 30 is coupled to the packet data network 34, which is connected to various endpoints, such as computers and network telephones (not shown) (a network telephone is a telephone that is fitted with a network interface card for communications over packet data networks). Examples of packet-switched communications include web browsing, electronic mail, text chat sessions, file transfers, interactive game sessions, packet-switched voice sessions, push-to-talk sessions, and so forth.

[0020] The wireless communications network thus provides two different types of communications: circuit-switched communications and packet-switched communications.

- 5 -

Circuit-switched communications are routed through the MSC 24, while packet-switched communications are routed through the PDSN 30. In circuit-switched communications, a dedicated end-to-end circuit or channel is established for the duration of a call session.

However, packet-switched communications utilize a connectionless intranetwork layer, such as that defined by the Internet Protocol (IP). One version of IP, referred to as IPv4, is described in Request for Comments (RFC) 791, entitled "Internet Protocol," dated September 1981; and another version of IP, referred to as IPv6, is described in RFC 2460, "Internet Protocol, Version 6 (IPv6) Specification," dated December 1998. In packet-switched communications, packets or other units of data carry routing information (in the form of network addresses) that are used to route the packets or data units over one or more paths to a destination endpoint.

[0021] The mobile station 17A, 17B are capable of participating in packet-switched real-time, interactive communications sessions with each other, or with a node coupled to data network 34, through the wireless communications network (which includes the BTS 20A, 20B, BSC 22, and PDSN 30). The BSC 22 includes a packet based processing function, such as a Packet Control Function (PCF) for CDMA-based wireless access technology that directly interfaces (via A10 and A11 as defined by 3GPP2) with the PDSN. In another embodiment, the BSC 22 may include the packet SGSN and GGSN functions as defined by 3GPP to provide an interface to any packet based networks (such as Internet, intranet, etc.).

[0022] A "real-time, interactive communications session" refers to an exchange of data, such as audio and/or video and/or text, on a substantially real-time basis between two terminals. A session is substantially real-time if interaction is occurring between two terminals, in which a communication from one endpoint is followed relatively quickly by a response or another communication from the other endpoint, typically within seconds, for example. A real-time, interactive communications session is distinguished from exchanges of electronic mail (e-mail) between two terminals.

[0023] A "packet-switched real-time, interactive communications session" refers to a real-time, interactive communications session that is established over a packet-switched network, which can include a wired network and/or a wireless network. Examples of packet-switched real-time, interactive communications sessions include voice-over-IP call

- 6 -

sessions, press (push)-to-talk (PTT) sessions, text chat sessions, instant messaging sessions, and so forth.

[0024] The terms “press-to-talk” and “push-to-talk” are used interchangeably. A “PTT communications session” or “PTT session” refers to a communications session in which one user (the caller) is able to establish real-time, one-to-one or one-to-many voice communications (and/or other real-time communications such as video, or video and voice) by simply pressing a talk button or key on a mobile station. In response to activation of the talk button or key during a PTT session, a request is sent to a PTT server 36 that is coupled to the data network 34. The request sent by the mobile station in response to activation of the talk button or key is a floor control request that seeks the right or ability to talk in the PTT session.

[0025] Note that a PTT session involving two or more users is initially established by the use of call control signaling, such as Session Initiation Protocol (SIP) signaling or another type of call control signaling. SIP is described in RFC 2543, entitled “The Session Initiation Protocol (SIP) Specification,” dated March 1999. SIP defines control signaling for establishing multimedia sessions over packet-switched networks. Once the PTT session is established, any of the users involved can seek the right or ability to talk by activating the talk button or key on a respective PTT-enabled mobile station or other terminal. Note that a PTT endpoint (that is involved in a PTT session) can be any terminal, including wired (landline) terminals or wireless terminals. Voice-over-IP, text chat, and instant messaging sessions can also be established using SIP or another type of call control signaling.

[0026] Conventionally, during establishment of a packet-switched real-time, interactive communications session, such as any of the communications sessions listed above, various resources of the nodes in the communications network are allocated during the call setup procedure. A “call setup procedure” is a procedure performed in response to a “call setup request” or “call request.” A “call setup request” or “call request” refers to any request for setting up a communications session, such as a packet-switched real-time, interactive communications session. The allocation of resources during a call setup procedure can be time consuming, and adds to the delay experienced by the user. The delay caused by allocating certain resources during the call setup procedure may not be acceptable for certain

- 7 -

packet-switched, delay-sensitive applications, such as voice-over-IP applications, PTT applications, text chat or instant messaging applications, and so forth.

[0027] In accordance with some embodiments, at least some of the resources that are normally allocated during the call setup procedure are pre-allocated. Pre-allocation of resources can be performed during provisioning or registration of network nodes or mobile stations. The resources that are pre-allocated include software, hardware, and communications resources. Resources in the BTS 20A, 20B, BSC 22, PDSN 30, and/or PTT server 36 can be pre-allocated. A resource is said to be “normally” allocated during a call setup procedure if, according to current conventional communications protocols (wireless or wired protocols), the resource is allocated during the call setup procedure. A resource that is normally allocated during the call setup procedure is also referred to as a “call setup resource.”

[0028] Pre-allocation of resources can be performed in response to various events. For example, the pre-allocation of resources can occur in response to initial startup of the nodes of the communications network depicted in Fig. 1. Alternatively, or in addition, other events that can lead to pre-allocation of resources include initiation of a mobile station, or a group of mobile stations. For example, such other events include when a mobile station initially starts up in a wireless network, initially enters a wireless network, launches a real-time, interactive application (such as a PTT application, voice-over-IP application, or text chat/instant messaging application). In another embodiment, the pre-allocation of resources can be initiated by a specific user or group of users launching a real-time, interactive application (such as a PTT application, voice-over-IP application, or text-chat/instant-messaging application).

[0029] In the ensuing discussion, reference is made to pre-allocating resources for PTT sessions. However, note that the same or similar techniques can be applied to other forms of packet-switched real-time, interactive communications sessions.

[0030] As depicted in Fig. 1, a PTT session can be established between mobile stations 17A and 17B. Alternatively, a PTT session can be established between one of the mobile stations 17A, 17B, and a node coupled to the data network 34.

[0031] Each of the mobile stations 17A, 17B communicates with the PTT server 36 for establishing PTT sessions. The mobile station 17A communicates with the PTT server 36

- 8 -

through the BTS 20A, BSC 22, and PDSN 30. The mobile station 17B communicates with the PTT server 36 through the BTS 20B, BSC 22, and PDSN 30. For other forms of packet-switched, delay-sensitive communications, the PTT server 36 can be substituted with another type of server, such as a voice-over-IP server (e.g., SIP proxy server), a text chat or instant messaging server, a call session control function (CSCF) module (according to 3GPP or 3GPP2), and so forth.

[0032] Alternatively, the BSC 22 can be connected to the PTT server 36 without passing through the PDSN 30. The link between the BSC 22 and PTT server 36 in this alternative configuration is depicted by the dashed lines of Fig. 1.

[0033] The pre-allocation of resources in accordance with some embodiments also enables dedicated links to be established between the various nodes of the communications network depicted in Fig. 1. The dedicated link between the BTS 20A and the BSC 22 is a link 40, and the dedicated link between the BTS 20B and the BSC 22 is a link 42. The dedicated link between the BSC 22 and the PDSN 30 is a link 62, and the dedicated link between the PDSN 30 and PTT server 36 is a link 67. In the alternative embodiment where the BSC 22 is connected to the PTT server 36 without passing through the PDSN 30, the dedicated link between the BSC 22 and the PTT server 36 is a link 69.

[0034] Information 44 pertaining to pre-allocated resources of the BTS 20A is stored in a storage 46. Examples of resources (software, hardware, and/or communications resources) that can be pre-allocated in the BTS 20A include one or more of the following: channel elements (channels of trunks), modems (for modem communications, if any, between the BTS and BSC as well as between MS and BTS), RF (radio frequency) channels (RF channels to be used by a given mobile station or group of mobile stations), routing elements (router(s), if any, to be used for communications between the BTS and BSC), T1/E1 trunks, Ethernet links, static IP routes, BTS software elements, memory, processors, and other resources. Resources can be pre-allocated for a particular packet-switched, delay-sensitive application (e.g., PTT application, voice-over-IP application, text chat application, instant messaging application, etc.) that is executed in a given mobile station or group of mobile stations.

[0035] The link 40 between the BTS 20A and BSC 22 can be one of various types of links: a T1/E1 trunk, an Ethernet link, a static IP route, a dynamic IP route, and so forth. Pre-allocating a T1/E1 trunk refers to pre-assigning channel(s) of a T1/E1 trunk for

- 9 -

communications of a particular packet-switched, delay-sensitive application in a given mobile station or group of mobile stations. Pre-allocating an Ethernet link refers to pre-allocating a destination Ethernet address to which packets associated with the packet-switched, delay-sensitive application are to be routed. The destination Ethernet address contained in the information 44 relating to pre-allocated resources is an Ethernet address of the BSC 22.

[0036] A pre-allocated static IP route refers to an IP route having a desired quality of service (QoS) for communicating packets of a packet-switched, delay-sensitive application executable in a given mobile station or group of mobile stations. Pre-allocating IP routes (either static or dynamic) between nodes of the communications network involves allocating one or more IP routes having predetermined QoS for packet-switched, delay-sensitive applications. Packets associated with a particular packet-switched, delay-sensitive application are treated as higher-priority packets that are routed over the pre-allocated IP routes. A QoS framework, such as an IntServ (integrated services) or DiffServ (differentiated services) framework, can be used for routing the higher-priority packets over the pre-allocated IP routes. An example IntServ framework is provided by the Resource Reservation Protocol (RSVP), as described in RFC 2205, entitled "Resource Reservation Protocol (RSVP)," dated September 1997. A DiffServ framework is a reservation-less framework that provides differentiated classes of service for network traffic by classifying packets communicated over the network.

[0037] The BTS 20A includes a wireless interface 48 for communicating wireless signaling (e.g., RF signaling) with the mobile station 17A. Additionally, the BTS 20A includes a BSC interface 50 for communicating over the link 40 with the BSC 22. The BTS 20B contains the same elements as the BTS 20A.

[0038] The BSC 22 also includes a storage 52 for storing information used for pre-allocating resources in the BSC 22 for a particular packet-switched, delay-sensitive application in a mobile station or group of mobile stations. The information stored in the storage 52 includes user-related information 54, binding information 55, mobility information 56, and information 57 for other pre-allocated resources.

[0039] The user-related information 54 includes information such as the network access identifier (NAI) of a mobile station (user), an IP address of a mobile station (user), the quality

- 10 -

of service (QoS) assigned to the mobile station (user) for a packet-switched, delay-sensitive application, the grade of service (GoS) assigned to the mobile station (user) for a packet-switched, delay-sensitive application, and a pointer (*e.g.*, a binary-coded unique number) assigned to the mobile station (user) to enable the pre-allocation or reserving of resources for the particular mobile station (user). Such pointers are used by the network elements (*e.g.*, BTS, BSC, PDSN, PTT server, and any link connecting these nodes) to allow the mobile station or a group of mobile stations (associated with a specific pointer) to claim and use these pre-allocated resources (that are tied to a specific pointer) during the call setup procedure. Note that assigning resources to a mobile station is used interchangeably with assigning resources to a user.

[0040] Instead of user-related information pertaining to a single mobile station (user), the user-related information 54 can contain information pertaining to a group of mobile stations (users). For example, users can be divided into various levels, such as a first level, second level, and third level, which are associated with different bandwidth, channel, and/or circuit allocations for a packet-switched, delay-sensitive applications.

[0041] The binding information 55 establishes a relationship between the radio domain and the packet domain for a specific mobile station or a group of mobile stations. One example of this is to create a binding between a radio-related address of a mobile station, such as the international mobile subscriber identity (IMSI), mobile identification number (MIN), or mobile equipment ID (MEID), and the packet-related address of the mobile station, such as the IP address, NAI, and Ethernet address. As with the user-related information, the binding information 55 is pre-allocated during provisioning or registration of a mobile station, or at any other time prior to a call setup procedure in which the resources would normally be allocated.

[0042] For example, an IMSI of a mobile station can be mapped to the mobile station's NAI in such a way that the software resources needed for binding are all allocated during the provisioning or registration time, or at some other time prior to the call setup procedure. When the mobile station initiates a session, the mobile station does not need to have the network allocate the resources needed for binding the radio and packet domains.

[0043] The mobility information 56 is related to the mobility behavior of a mobile station or a group of mobile stations. As an example, the mobility information 56 can identify the

- 11 -

active cell sectors that are to be used for soft handoff service of a particular mobile station or group of mobile stations. Also, the mobility information 56 can specify transcoders to be used for transforming air interface specific vocoding (such as EVRC) to landline-based vocoding (such as PCM, G-711, and so forth). The mobility information 56 also identifies
5 transport resources between the BTS, BSC, and MSC.

[0044] The BSC 22 includes a BTS interface 58 for communicating over the link to the BTS 20A. A similar BTS interface (not shown) is provided for communicating with the BTS 20B. Additionally, the BSC 22 includes an A10/A11 interface 60 for communicating over link 62 with the PDSN 30. An A11 interface defined between the BSC 22 and the PDSN 30 carries
10 control signaling between these two nodes. The A10 interface carries user traffic between the BSC 22 and PDSN 30.

[0045] Information 66 pertaining to pre-allocated resources of the PDSN 30 is stored in a storage 68. The PDSN 30 also includes an A10/A11 interface 64 for communicating with the BSC 22. The pre-allocated resources of the PDSN 30 include routing elements, A10/A11
15 interfaces, T1/E1 trunks, PDSN software elements, memory, processors, and other resources for a particular packet-switched, delay-sensitive application. The PTT server 36 also contains a storage 74 that stores information 76 pertaining to pre-allocated resources of the PTT server 36. The pre-allocated resources of the PTT server 36 include routing elements, transcoders, T1/E1 trunks, PTT server software elements, memory, processors, and other resources for a
20 mobile station or group of mobile stations.

[0046] As further depicted in Fig. 1, a management system 70 includes a network management module 72. In one embodiment, the management module 72 is a software module executable on processor(s) in the management system 70. The management system 70 is capable of communicating with the BTS 20A, 20B, BSC 22, PDSN 30, and PTT server
25 36 to provision these nodes, which includes the pre-allocation of resources in these nodes. Under control of the network management module 72, the management system 70 responds to predetermined events by sending requests to the various nodes of the communications network to pre-allocate resources.

[0047] Fig. 2 illustrates an alternative embodiment of a communications network. In the
30 embodiment of Fig. 2, mobile stations 17A, 17B communicate wirelessly with respective BTS 100A, 100B. Unlike the arrangement of Fig. 1, the BTS 100A, 100B are capable of

- 12 -

communicating over links 104, 106 with a PTT server 102. Consequently, the BTS 100A, 100B do not need to communicate through the BSC 108 with the PTT server 102. The BSC 108 also communicates with a PDSN 110 that is coupled to the data network 34.

[0048] A management system 112, on which a management module 114 is executable, also performs provisioning (including pre-allocating resources) of nodes of the communications network of Fig. 2, including nodes BTS 100A, 100B, BSC 108, PDSN 110, and PTT server 102. The management system 112 is similar to the management system 70 of Fig. 1.

[0049] As further depicted in Fig. 2, the PTT server 102 is coupled to an AAA (Authentication, Authorization, and Accounting) server 116. The AAA server 116 provides authentication, authorization, and accounting services for network nodes. AAA is described in RFC 2977, entitled "Mobile IP Authentication, Authorization, and Accounting Requirements," dated October 2000.

[0050] The AAA server 116 is also accessible by the other nodes of the network, including BTS 100A, BSC 108, and PDSN 110. Also, although not depicted in Fig. 1, an AAA server can also be provided in the communications network of Fig. 1.

[0051] Fig. 3 is a message flow diagram of a process according to an embodiment. Initially, network nodes (including the management system, BTS, BSC, PDSN, and PTT server) are started up and brought into operational mode (at 200). At this point, the management system is able to provision the other network nodes, including the BTS, BSC, PDSN, and PTT server. Part of this provisioning includes the pre-allocating of resources for a packet-switched, delay-sensitive application in a mobile station or group of mobile stations.

[0052] The management system performs pre-allocation of resources in response to predetermined events, such as an initial startup of nodes of the communications network (at 200), initiation of a mobile station (such as when a mobile station first starts up, first enters a wireless network, or launches a packet-switched, delay-sensitive application). In response to one of these predetermined events, the management system sends (at 202) to the BTS a request to pre-allocate BTS resources. The BTS performs resource pre-allocation in response to this request and stores the information 44 (Fig. 1) relating to the pre-allocated resources. The BTS then sends (at 204) an acknowledgment that the BTS resources have been reserved.

- 13 -

[0053] Similarly, the management system sends (at 206) to the BSC a request to pre-allocate BSC resources. In response to this request, the BSC pre-allocates resources and stores information pertaining to such pre-allocated resources in the storage 52 (Fig. 1). The BSC then sends (at 208) an acknowledgment that the BSC resources have been reserved.

5 [0054] The management system also sends (at 210) to the PDSN a request to pre-allocate PDSN resources. In response, the PDSN pre-allocates resources, and stores information 66 (Fig. 1) pertaining to such pre-allocated resources in its storage 68 (Fig. 1). The PDSN then sends (at 212) an acknowledgment that the PDSN resources have been reserved. Next, the management system sends (at 214) to the PTT server a request to pre-allocate PTT server
10 resources. In response, the PTT server performs pre-allocation of resources and stores information 76 pertaining to such pre-allocated resources in its storage 74 (Fig. 1). The PTT server then sends (at 216) an acknowledgment that the PTT server resources have been reserved.

[0055] At this point, pre-allocation of resources for packet-switched, delay-sensitive
15 application(s) has been performed to enable faster setup of packet-switched real-time, interactive communications sessions.

[0056] In the Fig. 3 example, the mobile station sends (at 218) an origination message (call setup request) to establish a PTT call session. The origination message sent at 218 can include a PTT service option (SO) information element to identify the request as being
20 associated with a PTT session. The service option information element is used by the BTS to determine that the message is related to a PTT session, so that the BTS can route the origination message accordingly.

[0057] Next, the BTS identifies (at 220) the PTT user, based on information contained in the origination message. Instead of the BTS performing the identification of the PTT user, the
25 BSC can perform the identifying act. The BTS/BSC can identify the PTT user by looking at the service option information element or by accessing a subscriber database. The BTS/BSC also performs authentication of the user (described further below) as part of the identifying act.

[0058] The BTS sends (at 222) a PTT request to the BSC, using pre-allocated resources
30 (including dedicated channels such as T1/E1 trunks, Ethernet links, or static IP routes). The

- 14 -

BSC forwards (at 224) the call setup message (such as an A11-Registration request) to the PDSN, again using pre-allocated resources (such as a pre-allocated A10/A11 interface).

[0059] The PDSN sets up the call by sending a PTT request (at 226) to the PTT server (using pre-allocated resources). The PTT server responds (at 228) with a PTT response. The PDSN
5 sends a PTT response through the same (or other) pre-allocated resources to the BSC. The BSC in turn sends (at 232) a PTT response to the BTS, again using pre-allocated resources. The BTS completes the PTT call setup and notifies (at 234) that the PTT call setup procedure has been completed.

[0060] After completion of call setup procedure, a user can obtain a right-to-speak by
10 activating a talk or speak button on the mobile station. The term "right-to-talk" or "ability-to-talk" refers to a particular terminal acquiring a token or other indication that a terminal is now able to transmit voice or other forms of real-time data to other terminals involved in the PTT call session. The mobile station initiates (at 236) this right-to-talk by sending a floor control request. The PTT server acknowledges (at 238) by granting the right-
15 to-talk.

[0061] By reserving or pre-allocating resources at the nodes of a communications network, such as the BTS, BSC, PDSN, and PTT server, call setup time can be reduced for specific packet-based, delay-sensitive applications, such as PTT applications, voice-over-IP applications, text chat applications, and instant messaging applications. Instead of allocating
20 these resources at call setup time, the resources are pre-allocated while provisioning the network, prior to call setup time.

[0062] As discussed above, in identifying the PTT user, authentication of the user is performed. Authentication can be performed by accessing an AAA server or by other techniques. Also, for quicker authentication, the BTS or BSC can maintain a cache of the
25 most recently authenticated users. In this manner, the BTS or BSC can access local storage to quickly authenticate a user. If a user cannot be found in the cache, then the BTS or BSC can access an external database.

[0063] For faster call setup, the authentication process can be performed in the background while the call setup procedure continues. Thus, the PTT request can be forwarded by the
30 BTS to the BSC and by the BSC to the PDSN, and so forth, without having to wait for completion of the authentication process. However, the authentication will have to be

- 15 -

completed before the PTT call setup complete notification can be sent (at 234) to the mobile station.

[0064] Instructions of the various software modules (e.g., software modules executed in BTS, BSC, PDSN, PTT server, management system) are loaded for execution on corresponding.

5 Processors include microprocessors, microcontrollers, processor modules or subsystems (including one or more microprocessors or microcontrollers), or other control or computing devices. As used here, a “controller” refers to hardware, software, or a combination thereof. A “controller” can refer to a single component or to plural components (whether software or hardware).

10 [0065] Data and instructions (of the software) are stored in respective storage devices, which are implemented as one or more machine-readable storage media. The storage media include different forms of memory including semiconductor memory devices such as dynamic or static random access memories (DRAMs or SRAMs), erasable and programmable read-only memories (EPROMs), electrically erasable and programmable read-only memories
15 (EEPROMs) and flash memories; magnetic disks such as fixed, floppy and removable disks; other magnetic media including tape; and optical media such as compact disks (CDs) or digital video disks (DVDs).

[0066] The instructions of the software are loaded or transported to each entity in one of many different ways. For example, code segments including instructions stored on floppy
20 disks, CD or DVD media, a hard disk, or transported through a network interface card, modem, or other interface device are loaded into the entity and executed as corresponding software routines or modules. In the loading or transport process, data signals that are embodied in carrier waves (transmitted over telephone lines, network lines, wireless links, cables, and the like) communicate the code segments, including instructions, to the entity.
25 Such carrier waves are in the form of electrical, optical, acoustical, electromagnetic, or other types of signals.

[0067] While some embodiments have been disclosed with respect to a limited number of embodiments, those skilled in the art will appreciate numerous modifications and variations there from. It is intended that the appended claims cover such modifications and variations as
30 fall within the true spirit and scope of the invention.